**Architecture Document for**

**Stock Market Prediction**

.........

TMP-AD*-*AD-5-002

Table of Contents

[1 Introduction – PART-A 1](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661441)

[1.1 Purpose 1](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661442)

[1.2 Scope 1](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661443)

[1.3 Definitions, Acronyms and Abbreviations 1](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661444)

[1.4 References 2](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661445)

[2 Architectural Goals and Constraints - – PART-B 3](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661446)

[2.1 Reusability 3](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661447)

[2.2 Scalability](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661448) 3

[2.3 Customizability](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661449) 3

[2.4 Extendibility](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661450) 3

[2.5 Use of Existing Business Logic](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661451) 3

[2.6 Time to Market](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661452) 3

[2.7 Portability](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661453) 3

[3 Productization Assessment – PART-B 4](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661455)

[3.1 Re-Usable Components 4](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661456)

[3.2 Analyze Architectural Frameworks in Repository 4](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661457)

[4 System Architecture – PART-A/B 5](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661459)

[4.1 Overview - – PART-A 5](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661460)

[4.2 Logical/Functional View - – PART-A 7](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661461)

[4.3 Use Case View – PART-A 8](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661462)

[4.4 Deployment View – PART-B 10](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661463)

[5 Alternative Solutions Considered – PART-B 10](file:///D:\JATAYU\Jatayu%20-%20format.docx#_Toc57661464)

**Document Revisions**

# 1 Introduction

## Purpose

Predicting how stock market will move is one of the most challenging issues due to many factors that involved rates ,politics and economics growth that make the stock market volatile.

The main aim is to identify the sentiments of investors. It is usually difficult as there must be rigorous analysis of national and international events. It is very important for an investor to know the current price and get a very close estimation of the future price.

## Scope

Predicting the performance of a stock market is tough as it takes into account various factors. The main aim is to identify the sentiments of investors. Sometime the stock market does well even when the economy is falling because there are various reasons for the profit or loss of a share.

Stock prices can be analyzed by the Investors as follows..

1. Get Real-Time stock price
2. View historical stock price
3. View the trend in market index
4. Analyze the trend in stock price
5. View amount of foreign handling

## Definitions, Acronyms and Abbreviations

CNN --- Convolutional Neural Networks

LSTM --- Long Short Term Memory

MLP --- Multilayer Perceptron

ARIMA --- Autoregressive Integrated Moving Average Model

ANN --- Artificial Neural Network

RFE --- Recursive Feature Elimination

PCA --- Principal Component Analysis

## References

**[1]** R. Vanaga and B. Sloka, “Financial and capital market commission financing: aspects and challenges,” *Journal of Logistics, Informatics and Service Science*, vol. 7

pp. 17–30, 2020.

**[2]** L. Zhang and H. Kim, “-e influence of financial service characteristics on use intention through customer satisfaction with mobile fintech,” *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 82–94, 2020.

**[3]** L. Badea, V. Ionescu, and A.-A. Guzun, “What is the causal relationship between stox europe 600 sectors But between large firms and small firms” *Economic Computation And Economic Cybernetics Studies And Research*, vol. 53, no. 3, pp. 5–20, 2019.

**[4]** J. Sousa, J. Montevechi, and R. Miranda, “Economic lot-size using machine learning,parallelism, metaheuristic and simulation,” *Journal of Logistics, Informatics and Service Science*, vol. 18, no. 2, pp. 205–216, 2019.

**[5]**  Dicle Mehmet F., John Levendis

**Importing financial data.** The Stata Journal, 14 (4) (2011), pp. 620-626

**[6]** Hossain M A, Karim R, Thulasiram R K, Bruce N D B, Wang

Y. Hybrid Deep Learning Model for Stock Price Prediction. IEEE Symposium Series on Computational Intelligence (SSCI), 2018, pp. 1821.

**[7]** Yang B, Gong Z J, Yang W. Stock Market Index Prediction Using Deep Neural Network Ensemble. 36th Chinese Control Conference (CCC), 2017, pp. 262

**[8]** Luca D P, Honchar O. Recurrent Neural Networks Approach to the Financial Forecast of Google Assets. International Journal of Mathematics and Computers in simulation, 2017, vol. 11, pp. 713.

***2 Architecture Goals And Constraints:***

***2.1 Reusability***

We can reuse the algorithm of this project for predicting the stock market price. We have used the module tensorflow and keras for the prediction values.Matpotlib is used for plotting a graph.Therefore the Modules of this project will be reusable,secure and reliable.

***2.2 Scalability***

The algorithm of this project can be adapted to any of the dataset.We can predict the stock for a very large amount of data.If we increase the Volume of Dataset,it will not affect the performance of the Prediction.

***2.3 Customizability***

We can customize the design of the User Interface as we want.We can add up an extra features as well as customizing the interface.

***2.4 Extendibility***

We can extend the prediction of the dataset for future growth.It allows us to bring a innovative solutions without any risking the ongoing operations.We can easily add the new modules and features in the future.

***2.5 Use of Existing Business Logic***

We can reuse the dataset of the any existing business data models such as DLL.We can also implement other concept of idea to predict the stock data using the same User Interface.

***2.6 Time to Market***

As a image predecessor,CNN algorithm in deep learning will be the good choice to train and test the model for accurate classification.Here we used this algorithm to get a highly predicted true dataset.

***2.7 Portability***

Our product will be a simple web app that can be able to run at any of the Operating system(OS).There will be no more need of any high system requirements to run the application. It also can run at the different Browsers like Chrome,Microsoft Edge,Mozilla Firefox etc…

***3 Productization Assessment***

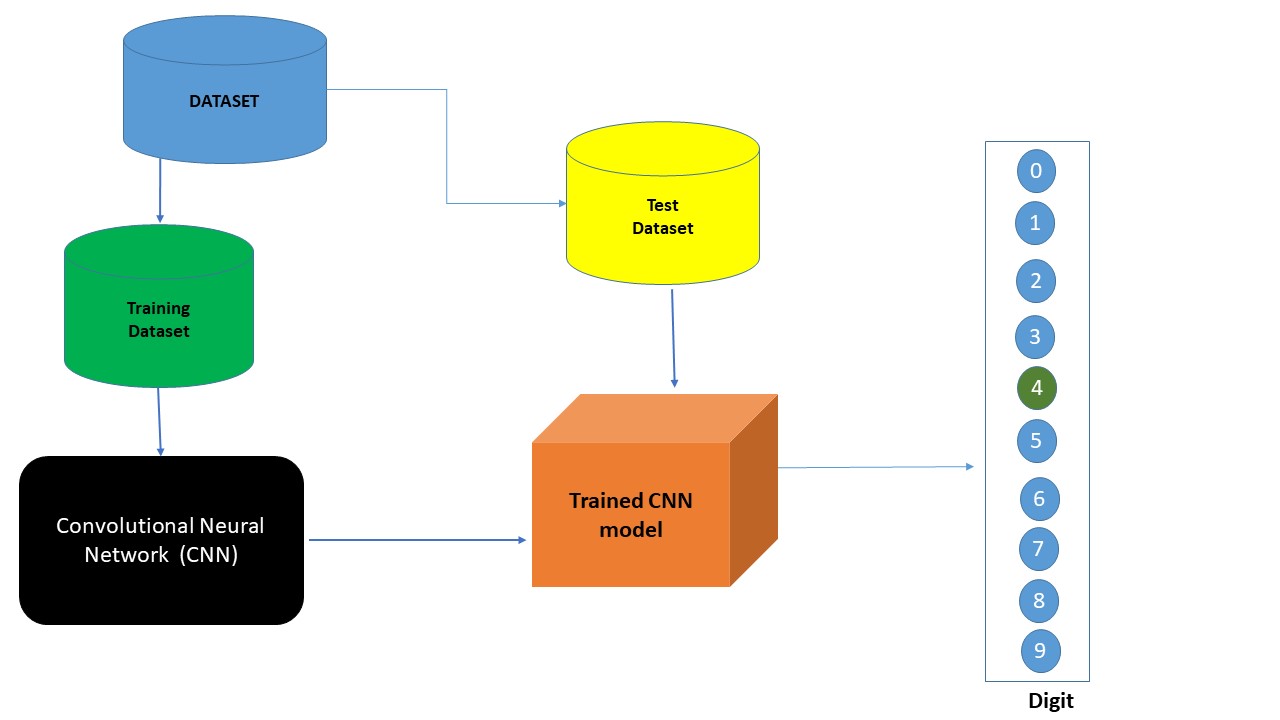
***3.1 Re-Usuable component***

We can reuse the whole part of the User Interface or algorithm or particular part in a UI for any further development process.It can be reliable for development purpose.

For an instance,we can say that if we need to change a algorithm of a model to any other algorithm we can able to achieve the process.We have used the module tensorflow and keras for the prediction values.Matpotlib is used for plotting a graph

***3.2 Analyse Architectural Frameworks in Repository***

We can use a high amount of dataset at time. Convolutional Neural Network algorithm can predict the huge dataset from a repository or from the database.



# 4 System Architecture

CNN has the characteristic of paying attention to the most obvious features in the line of sight, so it is widely used in feature engineering. LSTM has the characteristic of expanding according to the sequence of time, and it is widely used in time series. According to the characteristics of CNN and LSTM, a stock forecasting model based on CNN-LSTM is established.

We focus on the short-term price trend prediction. Currently, we only have the raw data with no labels. So, the very first step is to label the data.

If the price trend goes up, we mark it as 1 or mark as 0 in the opposite case. To be more specifed, we use the indices from the indices of *n*−*1th* day to predict the price trend of the *nth* day.

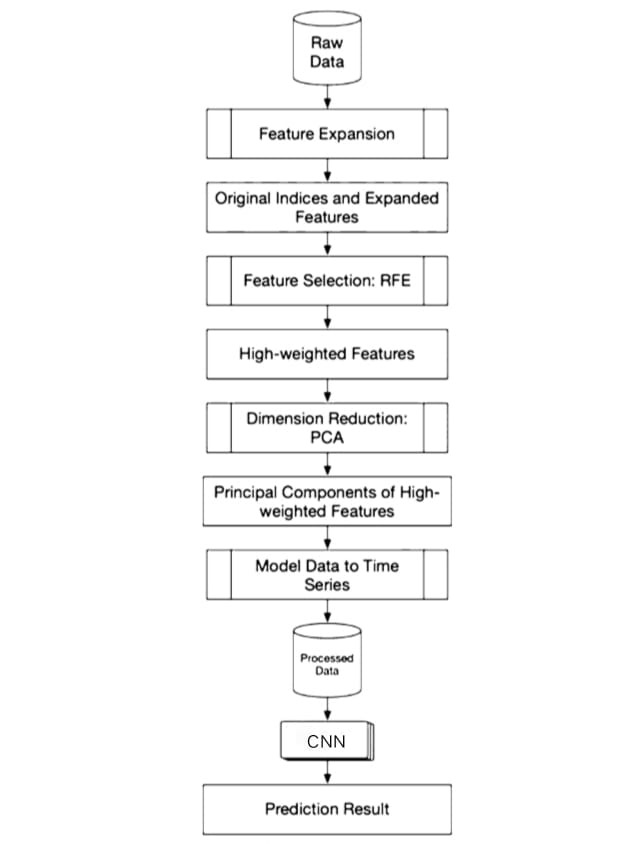
Machine learning algorithms are either supervised or unsupervised. In Supervised learning, labelled input data is trained and algorithm is applied. Classification and regression are types of supervised learning. It has a higher controlled environment.

## Overview -

The high-level architecture of our proposed solution could be separated into three parts. First is the feature selection part, to guarantee the selected features are highly effective. Second, we look into the data and perform the dimensionality reduction.

The last part, which is the main contribution of our work is to build a prediction model of target stocks. There are ways to classify different categories of stocks. Some investors prefer long term investments, while others show more interest in short-term investments. It is common to see the stock-related reports showing an average performance, while the stock price is increasing drastically; this is one of the phenomena that indicate the stock price prediction has no fixed rules, thus finding effective features before training a model on data is necessary.

**Overall Architecture of the Application:**



## Logical/Functional View

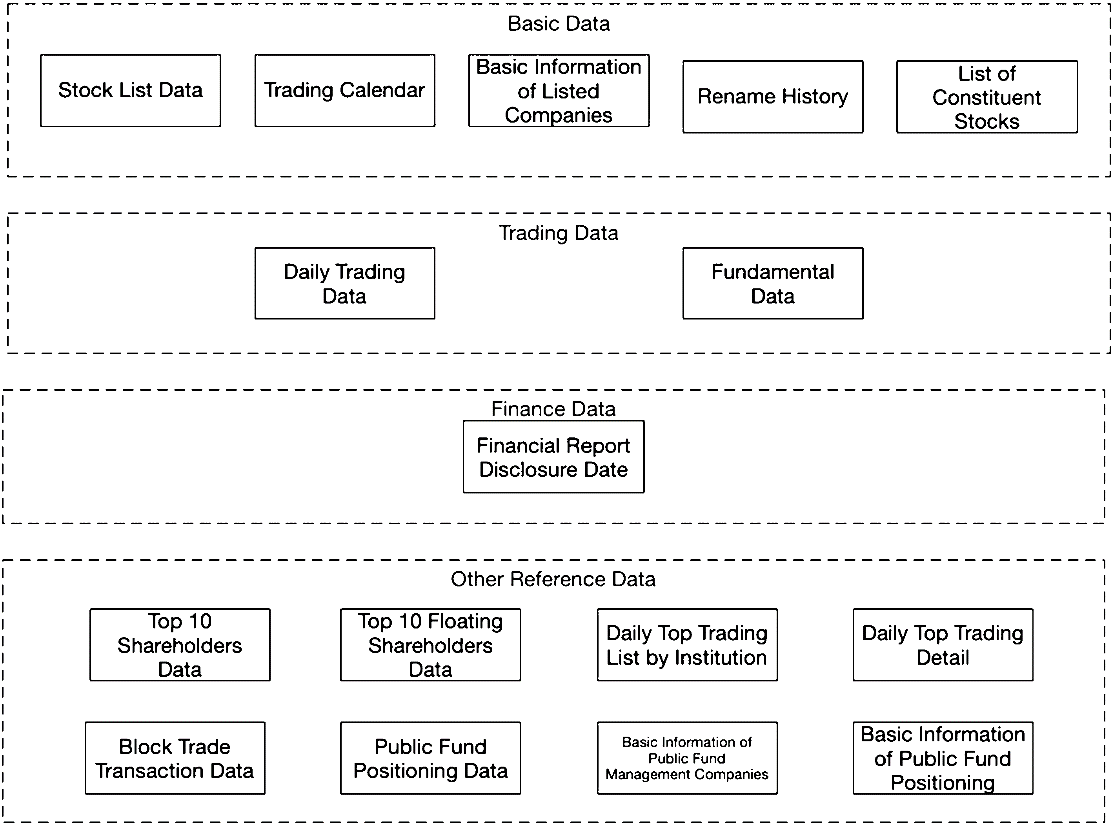
A **Convolutional Neural Network** (**CNN**) is a class of [deep neural networks](https://en.wikipedia.org/wiki/Deep_neural_network), most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution.

Each convolution layer contains a plurality of convolution kernels, and its calculation formula is shown in formula ([1](https://www.hindawi.com/journals/complexity/2020/6622927/#EEq1)). After the convolution operation of the convolution layer, the features of the data are extracted, but the extracted feature dimensions are very high, so in order to solve this problem and reduce the cost of training the network, a pooling layer is added after the convolution layer to reduce the feature dimension:

Li= tanh(xt \*kt +bt) ------------> 1

This section details the data that was extracted from the public data sources, and the final dataset that was prepared. Stock market-related data are diverse, so we first compared the related works from the survey of financial research works in stock market data analysis to specify the data collection directions.

**Block View of The Data:**

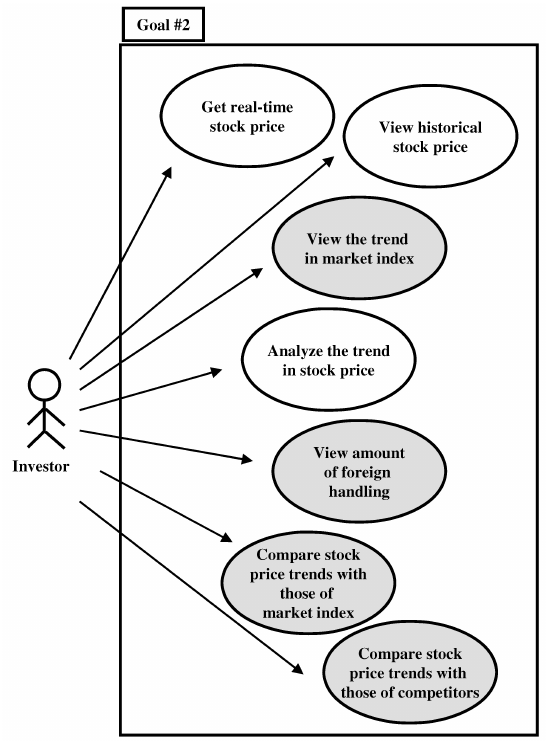


**4.3** **Use Case View**

**Investor:**

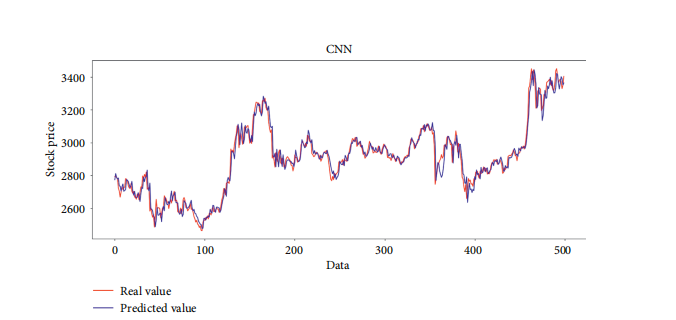
Real-time data is extensively used in many trading applications to create charts that allow traders to monitor moves or directions in a stock's price. Here the real time datasets are used in process.

Investor views the real time stock price and historical stock price and compare the stock price trend with market index and competitors.



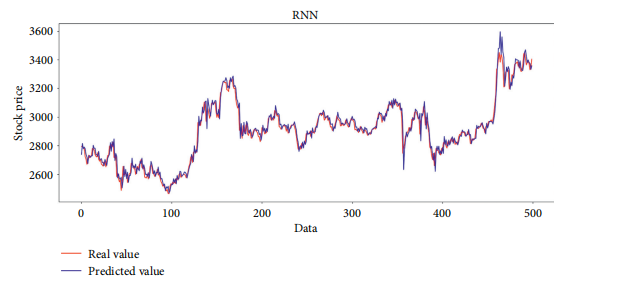
**Data Prediction**

Among the six forecasting methods, the broken line fitting degree of real value and predicted value is CNN, RNN, and MLP. CNN has the highest degree of broken line fitting which almost coincides with each other, and MLP has the lowest degree of broken line fitting.



Comparison of the predicted value and the real value for CNN.

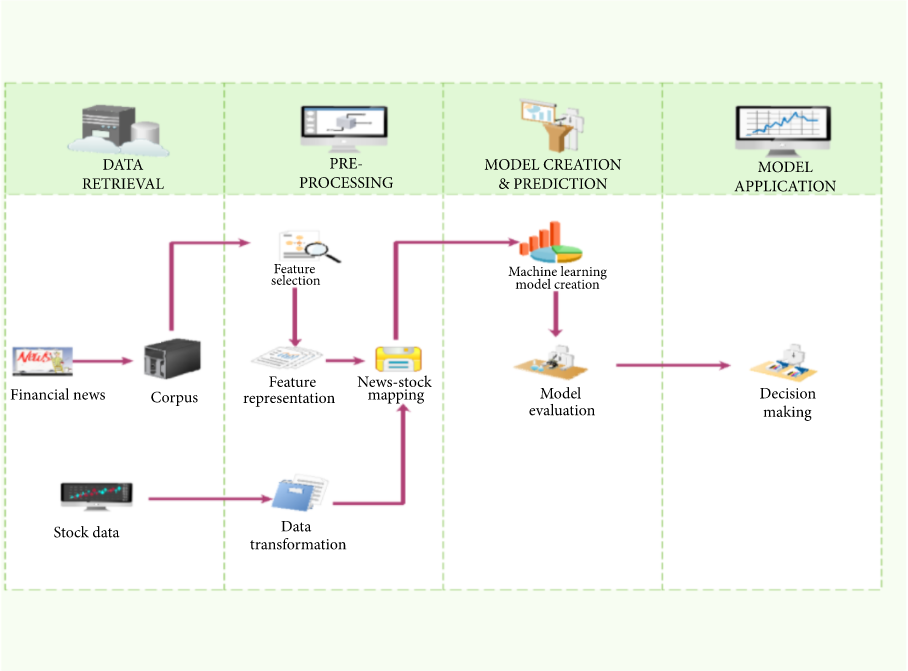
According to the predicted value and real value of each method, the evaluation index of each method can be calculated



* 1. ***Deployment View***

While deploying the web app on any system will be going to me much easier for our product.Because we had implemented a simple User Interface which takes the dataset.

Our application takes the dataset which is given as a input,then the data will be pre-processed to eliminate the unwanted data.

Then the model data will be trained and tested for the evaluation of the model.Finally the decision of the prediction for the dataset will be made.

1. ***Alternative Solutions Considered***

* To predict the market, mostly use either technical or fundamental analysis.  Technical analysis focuses on analyzing the direction of prices to predict future prices, while fundamental analysis depends on analyzing unstructured textual information like financial news and earning reports.
* Earlier classical regression methods such as linear regression, polynomial regression, etc. were used to predict stock trends.
* Nowadays, Support Vector Machines (SVM) and Artificial Neural Networks (ANN) are widely used for the prediction of stock price movements. Every algorithm has its way of learning patterns and then predicting.
* The proposed CNN achieved relatively higher prediction accuracy of 68.6%, while the ANN, SVM, and KNN algorithms obtained prediction accuracies of 73.5%, 67.9%, and 65.9%